# A Balanced Diagnostic System Compatible with a Barotropic Prognostic Model

MICHAEL GHIL

*Courant Institute of Mathematical Sciences, New York University, New York, N. Y. 10012, and NASA Institute for Space Studies, Goddard Space Flight Center, New York, N. Y. 10025*

BORIS SHKOLLER[1] AND VICTOR YANGARBER

*General Telephone and Electronics, Information Systems, Inc., New York, N. Y. 10025*

(Manuscript received 14 December 1976, in revised form 13 June 1977)

ABSTRACT

We derive a system of diagnostic equations for the velocity field, or "wind laws," for a barotropic primitive-equation model of large-scale atmospheric flow. The derivation is mathematically exact and does not involve any physical assumptions, such as nondivergence or vanishing of derivatives of the divergence, which are not already present in the prognostic equations. Therefore, initial states computed by solving these diagnostic equations should be compatible with the type of motion described by the prognostic equations of the model, and should not generate initialization shocks when inserted into the prognostic model.

Based on the diagnostic system obtained, we are able to give precise meaning to the question whether the wind field is determined by the mass field and by its time history. The answer to this important question is affirmative, in the precise formulation we provide.

The diagnostic system corresponding to the chosen barotropic model is a generalization of the classical balance equation. The ellipticity condition for this system is derived and given a physical interpretation. Numerical solutions of the diagnostic system are exhibited, including cases in which the system is of mixed elliptic-hyperbolic type.

Such diagnostic systems can be obtained for other primitive equation models. They are valid for all atmospheric scales and regions for which the prognostic models from which they are derived hold. Some problems concerning the possibility of implementing such a system in operational numerical weather prediction are discussed.

## 1. Introduction

Efforts to solve the initialization problem have proceeded mainly along two lines: 1) *static* initialization, and 2) *dynamic* initialization (Bengtsson, 1975). Both are based on the theory of geostrophic adjustment and try to apply it in different ways. *Static initialization* attempts to derive diagnostic relations (i.e., relations free of time derivatives) among variables; these relations are compatible with prognostic systems the solutions of which do not contain components of the inertia-gravity wave (IGW) type. Initial states obtained from these relations should therefore have corresponded to a balanced state, or a state of the system after adjustment (Hinkelmann, 1951). It turned out, however, that using such initial states did not prevent the rapid growth of unwanted components of solutions to primitive-equation systems (Nitta and Hovermale, 1969; Morel *et al.*, 1971), which constitutes the *initialization shock*. This incompatibility effect could be explained by the fact that quasigeostrophic equations obtain from the primitive ones by a singular, rather than by a regular perturbation. The incompatibility effect has been studied recently in a filtered barotropic model by Blumen (1976a, b).

*Dynamic initialization* attempts to use the dispersive and dissipative properties of the numerical primitive-equation models themselves to reduce the amplitude of spurious IGW-type phenomena in the model, i.e., let the model simulate the geostrophic adjustment process in the atmosphere (Nitta and Hovermale, 1969). In such a procedure, the model is furnished additional data as they become available in order to compensate for the incompleteness of the initial data—the model is *updated* (Charney *et al.*, 1969). The hope is that, after a sufficient period of updating, the model's solution will closely approximate the state which it would have attained if given the proper set of complete initial data.

The difficulties encountered with this technique and variations thereof (Morel *et al.*, 1971; Williamson and Kasahara, 1971) were the following: 1) decrease of the initial rms error in the solution to a nonzero asymptotic value, 2) achievement of this decrease within periods not shorter than 2 days, and 3) slower decrease or actual

[1] Current affiliation: National Center for Atmospheric Research, Boulder, Colo. 80307.

increase of the rms error when real, rather than simulated, data were used. These difficulties are connected with the nonlinearity of the system and with the change of the solution during the updating process.

We propose to overcome these difficulties by reducing the updating interval to zero length, i.e., by obtaining the initial state as the solution of a generalized diagnostic system which includes time derivatives of those variables that are assumed to be known completely and accurately; in mid-latitudes it is reasonable to assume that this is the case for the mass field. This system is derived from the corresponding prognostic system by exact mathematical manipulations, without neglecting any terms which may be small, but have an entirely non-negligible effect on the behavior of the solution to the prognostic system. The compatibility of the diagnostic system with the prognostic system should prevent spurious growth of IGW-type phenomena (the initialization shock).

Such a diagnostic system was derived first for a *linear barotropic* primitive-equation model in Ghil (1973, 1975a, b); a complete analysis of this linear diagnostic system and of the effect of using its solution as an initial state for the corresponding linear prognostic system was given. Furthermore, a diagnostic system compatible with a *nonlinear baroclinic* primitive-equation model in pressure coordinates was derived in Ghil (1975b). Similar derivations can be given in $z$ coordinates and in $\sigma$ coordinates.

We shall present here a *nonlinear barotropic* case of such a compatible diagnostic system, or balance system. It is of interest, however, to point out that the diagnostic system we present has exactly the same form as the one that obtains for a baroclinic primitive-equation model in isentropic coordinates (Ghil, 1975b).

The article is organized in the following way: Section 2 contains the derivation of the diagnostic system compatible with the nonlinear shallow-fluid equations. In it the ellipticity condition for this system is derived and given a physical interpretation. Connections with the classical balance equation and its ellipticity condition are pointed out. Section 3 presents the derivation of a second-order system which is equivalent to the first-order system in Section 2, but which is easier to solve numerically. It is shown that the domains of ellipticity of the two systems are closely related. Section 4 describes the numerical method used to solve the second-order system of the previous section, and contains the results for a number of test cases. The test cases include solutions for which the system is of mixed elliptic-hyperbolic type, as well as cases with non-zero divergence. Section 5 contains the summary and conclusions, as well as some comments on the possibility of applying the approach we present and the numerical methods we developed in an operational context.

## 2. A barotropic balance system

We consider the nonlinear shallow-fluid model governed by the equations

$$u_t+uu_x+vu_y+\phi_x-fv=0, \tag{1a}$$

$$v_t+uv_x+vv_y+\phi_y+fu=0, \tag{1b}$$

$$\phi_t+u\phi_x+v\phi_y+\phi(u_x+v_y)=0. \tag{1c}$$

Here $u$, $v$ are the velocity components in the $x$, $y$ directions, respectively, $f$ is the Coriolis parameter and $\phi$ is the geopotential of the free surface of the fluid.

We want to derive two equations for the velocity components $u$, $v$ in which their time derivatives do not appear, although $\phi_t$, $\phi_{tt}$ may be present. The assumption is that $\phi$ is known together with its time derivatives. The plausibility of this assumption will be discussed in the last section.

For convenience, we introduce the notations $\delta=u_x+v_y$ for the divergence, $d/dt=\partial/\partial t+u\partial/\partial x+v\partial/\partial y$ for the material derivative, and $\Phi=\log\phi$ ($\phi>0$). Differentiating (1a) with respect to $x$, (1b) with respect to $y$, and adding, we obtain the familiar divergence equation, which we write as

$$u_x^2+2u_yv_x+v_y^2+f(u_y-v_x)+f_yu+\nabla^2\phi=-\frac{d\delta}{dt}, \tag{2}$$

where $\nabla^2=\partial^2/\partial x^2+\partial^2/\partial y^2$ is the Laplacian. Eq. (1c) can be rewritten in our notation as $\delta=-(d\Phi/dt)$, and hence we also have

$$-\frac{d\delta}{dt}=\frac{d\Phi_t}{dt}+u\frac{d\Phi_x}{dt}+v\frac{d\Phi_y}{dt}+\Phi_x\frac{du}{dt}+\Phi_y\frac{dv}{dt}. \tag{3}$$

Expressing $du/dt$ and $dv/dt$ with the aid of (1a) and (1b), Eq. (3) becomes

$$-\frac{d\delta}{dt}=\Phi_x(fv-\phi_x)-\Phi_y(fu+\phi_y)+\frac{d\Phi_t}{dt}+u\frac{d\Phi_x}{dt}+v\frac{d\Phi_y}{dt}. \tag{3'}$$

Combining (3') with (2) we obtain a diagnostic equation of the desired type, to which we can add (1c) to obtain the balance system

$$u_x+v_y=-\Phi_xu-\Phi_yv-\Phi_t, \tag{4a}$$

$$u_x^2+2u_yv_x+v_y^2+f(u_y-v_x)=-f_yu+f(\Phi_xv-\Phi_yu)+\frac{d\Phi_t}{dt}+u\frac{d\Phi_x}{dt}+v\frac{d\Phi_y}{dt}-\frac{\phi_x^2+\phi_y^2}{\phi}-\nabla^2\phi. \tag{4b}$$

This is the diagnostic system compatible with the nonlinear shallow-fluid equations (1). Using as initial data for (1) the known $\phi$ and the $u$, $v$ obtained by solving (4) should yield a solution of (1) which is free of initialization shock. We further notice that no re-

strictions on scales of motion limit the validity of (4); this is an important distinction from the classical balance equation. In particular, system (4) reduces to the balance equation when $\delta=0=d\delta/dt$ (Ghil, 1975b).

In order to solve (4) it is important first to determine its mathematical type. This is equivalent to determining the existence and number of real characteristics of the system. If the system has no real characteristics at all it is *elliptic* and if it has the maximum possible number of distinct real characteristics it is *hyperbolic*; the maximum possible number of characteristics is simply the degree of the characteristic equation, which in the case at hand is two. If all characteristics are real, but some coincide, we shall call the system *parabolic.*

The generalization of the concept of *characteristics* to the nonlinear system (4) can proceed either from the concept applied to a simple nonlinear equation (Courant and Hilbert, 1962, pp. 418–421) or from that of the characteristics of a quasi-linear system (*ibid.*, pp. 170–173, 424–427). To derive the appropriate characteristic equation, we first rewrite (4) in the form

$$D(\mathbf{V}_x,\mathbf{V}_y)=d(x,y,\mathbf{V}), \tag{5a}$$

$$E(\mathbf{V}_x,\mathbf{V}_y)=e(x,y,\mathbf{V}), \tag{5b}$$

where $\mathbf{V}$ is the velocity vector, $\mathbf{V}=(u,v)$, and $\mathbf{V}_x$, $\mathbf{V}_y$ are its first derivatives, $\mathbf{V}_x=(u_x,v_x)$, $\mathbf{V}_y=(u_y,v_y)$; $D$ and $E$ are the left-hand sides, and $d$, $e$, the corresponding right-hand sides of (4a) and (4b), respectively. Furthermore, we let

$$\Psi(x,y)=\text{constant}$$

be the equation of a characteristic of system (5) in implicit form.

In the notation introduced above, the *characteristic equation* of (5), i.e., the equation whose solutions are the characteristics of (5), is simply (Ghil, 1975b)

$$0=\det\left\{\begin{pmatrix}\mathrm{D}_{u_x} & \mathrm{E}_{u_x}\\ \mathrm{D}_{v_x} & \mathrm{E}_{v_x}\end{pmatrix}\Psi_x+\begin{pmatrix}\mathrm{D}_{u_y} & \mathrm{E}_{u_y}\\ \mathrm{D}_{v_y} & \mathrm{E}_{v_y}\end{pmatrix}\Psi_y\right\}. \tag{6}$$

After some computations, (6) can be written as

$$(f-2u_y)\Psi_x^2+2(u_x-v_y)\Psi_x\Psi_y+(f+2v_x)\Psi_y^2=0. \tag{6'}$$

That system (4) be elliptic is equivalent to (6′) not having real solutions $\Psi_x/\Psi_y$ (or $\Psi_y/\Psi_x$); that is, it is equivalent to the condition that the discriminant $\mathfrak{D}$

$$\mathfrak{D}=(f-2u_y)(f+2v_x)-(u_x-v_y)^2 \tag{7}$$

be positive. The *ellipticity condition* for (4) is therefore

$$\mathfrak{D}>0, \tag{8}$$

i.e., system (4) is elliptic if (8) holds and it is hyperbolic if the opposite inequality, $\mathfrak{D}<0$, holds. But, expanding and rearranging (7), we can put it into the form

$$\mathfrak{D}=f^2+d^2-2e, \tag{7'}$$

where we used

$$-2u_xv_y=(u_x^2+v_y^2)-(u_x+v_y)^2.$$

This finally yields the ellipticity condition for (4) as

$$f^2+d^2-2e>0, \tag{8'}$$

which generalizes the corresponding condition for the classical balance equation (Ghil, 1975b).

A further transformation of (7) yields an interesting physical interpretation of the ellipticity condition. Using (4a, b) to expand the right-hand side of (7′) and regrouping terms, we obtain

$$\mathfrak{D}=(f+v_x-u_y)^2-[(u_x-v_y)^2+(u_y+v_x)^2]. \tag{9}$$

In this form, we easily recognize that $\mathfrak{D}$ is just the difference between the square of the *absolute vorticity* and that of the *total deformation* of the flow field $\mathbf{V}$. Both of these quantities are invariant under transformations of coordinates and they play a role in synoptic diagnosis. They are discussed in a somewhat similar context by Petterssen (1953). We believe, however, that (9) and its physical interpretation are new.

Form (7′) is important, since it stresses that $\mathfrak{D}$ depends directly only on $u$ and $v$ themselves, and not on their derivatives. The situation is quite similar to that of the classical balance equation, to which (4) reduces when $\delta=d\delta/dt=0$. Indeed, the ellipticity condition for the classical balance equation also depends only on first and not on second derivatives of its solution, i.e., on velocity components and not on their derivatives; in both cases the highest derivatives of the problem are not present in the ellipticity condition. The dependence of $\mathfrak{D}$ on $u$ and $v$ only is very useful in working with real data and with numerical solutions, since a function can be approximated with greater accuracy than its derivatives, whether it be from observations or from computations.

The analogy with the classical balance equation and a scale analysis of (7′) leads us to expect that atmospheric data with $\mathfrak{D}<0$ might occur (Ghil, 1975b). Therefore, the solvability of (4) in regions containing hyperbolic ($\mathfrak{D}<0$) as well as elliptic ($\mathfrak{D}>0$) subdomains is crucial, both theoretically and practically.

Theoretically, we would like to know whether the mass field and its time history do indeed determine the wind field for fully nonlinear systems. The conjecture that this is so was put forward by Charney *et al.* (1969) to justify their updating approach to dynamic initialization and four-dimensional data assimilation. The partial success of updating did not provide conclusive evidence either to prove or to disprove this conjecture. Assuming that this theoretical question is settled in the affirmative, we would like to verify practically whether solving system (4) directly is a better approach to initialization than updating.

### 3. An equivalent second-order system

A semi-direct, rather than iterative, method to solve system (4) itself appears elsewhere (Ghil and Balgovind, 1977). In this article we shall present solutions of (4)

obtained by using a simple iterative numerical technique to solve a second-order system derived from and equivalent to (4). An approach similar to the one described here is being pursued independently by A. Sundström (personal communication, 1976).

The starting point of the following derivation was the observation that the classical balance equation

$$f\nabla^2\psi+2(\psi_{xx}\psi_{yy}-\psi_{xy}^2)+f_y\psi_y=\nabla^2\phi, \tag{10}$$

with $\psi$ the streamfunction, is a close approximation to a Poisson equation for small Rossby number. That is, when the Rossby number and hence the nonlinear terms of (10) are small, the left-hand side of the equation is dominated by the Laplacian $\nabla^2\psi$. But Poisson equations are readily solved by iterative methods. Hence it was natural to attempt to put (4) into a form similar to (10), without assuming nondivergence, $\delta\equiv 0$; it is nondivergence which entails the existence of the streamfunction $\psi$ appearing in (10), with $u=-\psi_y$, $v=\psi_x$.

A form similar to (10) is easily achieved by cross-differentiation of (4). Using the notation $d$, $e$ of (5) for the right-hand sides of (4a) and (4b), we immediately obtain

$$f\nabla^2 u+(u_x^2+2u_yv_x+v_y^2)_y=fd_x+e_y, \tag{11a}$$

$$f\nabla^2 v-(u_x^2+2u_yv_x+v_y^2)_x=fd_y-e_x, \tag{11b}$$

where we took $f=$ constant for simplicity. Not only is the Laplacian of $u$, $v$ put in evidence in this form, but (11a, b) are actually *quasi-linear* equations. That is, they are linear in the highest derivatives $u_{xx}$, $u_{xy}$, ..., $v_{yy}$, rather than fully nonlinear, as was (4b). This immediately suggests a simple iterative method for solving (11); we postpone the description of this numerical method for the next section.

First we have to ask in what sense systems (11) and (4) are equivalent. Clearly, solutions of (4) which are sufficiently differentiable will also be solutions of (11). Conversely, however, (11) might have some solutions which are not also solutions of (4). In particular, since (11) is of higher order than (4), it requires more boundary conditions to determine its solutions uniquely. A judicious choice of the supplementary boundary conditions should ensure that there is a one-to-one correspondence between solutions of (11) and of (4).

The next question that arises about system (11) concerns its type and the relationship of its type to that of (4). Some readers could find the following discussion rather technical and might want to proceed directly to its conclusion, which is given immediately after Eq. (21).

For convenience of notation, let (11) be written symbolically as

$$G(\mathbf{V}'',\mathbf{V}')=g(\mathbf{V}',\mathbf{V}), \tag{12a}$$

$$H(\mathbf{V}'',\mathbf{V}')=h(\mathbf{V}',\mathbf{V}), \tag{12b}$$

where $\mathbf{V}'=(\mathbf{V}_x,\mathbf{V}_y)$ denotes the first derivatives and $\mathbf{V}''=(\mathbf{V}_{xx},\mathbf{V}_{xy},\mathbf{V}_{yy})$ the second derivatives of $\mathbf{V}=(u,v)$. In this notation, the characteristic equation of (11) is

$$0=\det\left\{\begin{pmatrix}G_{u_{xx}} & H_{u_{xx}}\\ G_{v_{xx}} & H_{v_{xx}}\end{pmatrix}\Psi_x^2+\begin{pmatrix}G_{u_{xy}} & H_{u_{xy}}\\ G_{v_{xy}} & H_{v_{xy}}\end{pmatrix}\Psi_x\Psi_y+\begin{pmatrix}G_{u_{yy}} & H_{u_{yy}}\\ G_{v_{yy}} & H_{v_{yy}}\end{pmatrix}\Psi_y^2\right\}. \tag{13}$$

Written out explicitly, Eq. (13) is equivalent to

$$0=\det\begin{pmatrix}f\Psi_x^2+2u_x\Psi_x\Psi_y+(f+2v_x)\Psi_y^2 & 2(u_y\Psi_x+v_y\Psi_y)\Psi_y\\ -2(u_x\Psi_x+v_x\Psi_y)\Psi_x & (f-2u_y)\Psi_x^2-2v_y\Psi_x\Psi_y+f\Psi_y^2\end{pmatrix}, \tag{13'}$$

which becomes, after some algebraic manipulations,

$$0=(f-2u_y)\Psi_x^4+(f+2v_x)\Psi_y^4+2(f+v_x-u_y)\Psi_x^2\Psi_y^2+2(u_x-v_y)\Psi_x\Psi_y(\Psi_x^2+\Psi_y^2). \tag{14}$$

The connection between (14) and (6′) is not immediately apparent. To put a connection in evidence we modify (11) slightly by using (4a) to eliminate the mixed derivatives $u_{xy}$, $v_{xy}$; i.e., we substitute into (11) the expressions

$$u_{xy}=d_y-v_{yy}, \quad v_{xy}=d_x-u_{xx}. \tag{15}$$

After substituting (15), system (11) becomes

$$Au_{xx}+Bu_{yy}+Cv_{yy}=e_y+Ad_x-2u_xd_y, \tag{16a}$$

$$Cu_{xx}+Av_{xx}+Bv_{yy}=-e_x+2v_yd_x+Bd_y, \tag{16b}$$

where we introduced the notations

$$A=f-2u_y, \quad B=f+2v_x, \quad C=2(v_y-u_x). \tag{17}$$

We now let the abbreviated form of (16) be

$$G_1(\mathbf{V}'',\mathbf{V}')=g_1(\mathbf{V}',\mathbf{V}), \tag{18a}$$

$$H_1(\mathbf{V}'',\mathbf{V}')=h_1(\mathbf{V}',\mathbf{V}), \tag{18b}$$

in the same way in which (12) stands conveniently for (11) and system (5) stands for system (4). The characteristic equation of (16) is (13), with $G$ and $H$ replaced by $G_1$ and $H_1$. Carrying out the operations of partial differentiation, we obtain

$$0=\det\begin{pmatrix}A\Psi_x^2+B\Psi_y^2 & C\Psi_y^2\\ C\Psi_x^2 & A\Psi_x^2+B\Psi_y^2\end{pmatrix}=A^2\Psi_x^4+(2AB-C^2)\Psi_x^2\Psi_y^2+B^2\Psi_y^4. \tag{19}$$

Clearly, Eq. (19) has four solutions, which are in general complex. These solutions come in pairs corresponding to the choice of square root for $\xi=\Psi_x^2/\Psi_y^2$ (or $1/\xi=\Psi_y^2/\Psi_x^2$). Therefore the solutions will be real and distinct whenever $\xi>0$, and they will be complex with nonzero imaginary part whenever $\xi<0$ or $\xi$ is not real. A necessary and sufficient condition that system (16) be elliptic is therefore that the discriminant $\mathcal{E}$ of (19), as an equation for $\xi$ (or $1/\xi$), be positive, i.e.,

$$0<\mathcal{E}\equiv 4A^2B^2-(2AB-C^2)^2 =C^2(4AB-C^2), \tag{20}$$

and that $\xi$ be positive itself. Recalling the definitions (17) of $A$, $B$, $C$, and the expressions (7,7′) for $\mathfrak{D}$, we obtain

$$\mathcal{E}=C^2\mathfrak{D}. \tag{21}$$

Thus, finally, Eq. (16) is elliptic whenever (4) is, except for parabolicity along the lines $C^2\equiv 4(u_x-v_y)^2=0$, and for hyperbolicity where $\xi$ is real, but negative. To summarize, we can say that the domain of ellipticity of (4) includes that of (16), and possibly exceeds it; we are not able at this point to make a similar clear-cut statement for (11).

Having partially clarified the relationship between systems (11) and (16) on the one hand, and system (4) on the other, we proceed to describe our test computations with system (11). Before doing so, we would like to mention two previous attempts at integrating diagnostic equations of mixed type numerically, those of Fjørtoft (1962) and Miyakoda and Moyer (1968). We hope that our contribution is a further step in the direction of solving this difficult problem.

## 4. Numerical method and results

### *a. The method*

We wish to solve the second-order system (11) in a rectangle $0\leqslant x, y\leqslant L$, $L=6\times 10^6$m, which is typical of a mid-latitude zonal channel, and we take $f=$constant $=10^{-4}$ s$^{-1}$. The boundary conditions we chose are to prescribe $u$ and $v$ on all sides of the rectangle. These boundary conditions are clearly appropriate for the elliptic linear problem corresponding to (11). The numerical experiments on which we report seem to indicate that they are suitable for the nonlinear problem as well, at least as long as the hyperbolic subdomains do not become too large.

To discretize (11) we introduce a grid with uniform mesh spacing $\Delta x=\Delta y=\Delta$, $L/\Delta=M$, and the usual notation $\mathbf{V}_{ij}=\mathbf{V}(i\Delta,j\Delta)$. Derivatives are replaced by centered differences, e.g.,

$$\mathbf{V}(x+\Delta,y)-\mathbf{V}(x-\Delta,y)\approx 2\Delta\mathbf{V}_x(x,y),$$

with similar approximations for $\mathbf{V}_y$, $\mathbf{V}_{xx}$, $\mathbf{V}_{xy}$ and $\mathbf{V}_{yy}$.

Introducing these expressions for the derivatives into (11) results in a system of *nonlinear algebraic equations* for $\mathbf{V}_{ij}$, $1\leqslant i, j\leqslant M-1$. It is worthwhile to notice that writing the equations at grid point $(i,j)$ only involves $\mathbf{V}_{ij}$ from the expressions of $\mathbf{V}_{xx}$ and $\mathbf{V}_{yy}$, and not from the first derivatives $\mathbf{V}_x$, $\mathbf{V}_y$ or from the mixed derivatives $\mathbf{V}_{xy}$. These equations are now rewritten for the purpose of solving them iteratively as

$$au_{ij}+bv_{ij}=q(u_{i+\alpha_1,j+\alpha_2}), \tag{22a}$$

$$cu_{ij}+dv_{ij}=r(u_{i+\beta_1,j+\beta_2}), \tag{22b}$$

where $\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2$ take the values $-1$, 0, 1, but always $|\alpha_1|+|\alpha_2|\neq 0\neq|\beta_1|+|\beta_2|$, i.e., $\mathbf{V}_{ij}$ is expressed as a function of the eight neighboring values, $\mathbf{V}_{i\pm1,j\pm1}$, $\mathbf{V}_{i,j\pm1}$, $\mathbf{V}_{i\pm1,j}$. The coefficients $a$, $b$, $c$, $d$ also depend on values of $\mathbf{V}$ at the neighboring points, but not on $\mathbf{V}_{ij}$; this is due to the quasi-linear character of system (11) and to the use of centered differences.

The *iterative procedure* now consists in choosing an initial guess $\mathbf{V}_{ij}^{(0)}$ and then defining recursively $\mathbf{V}_{ij}^{(n+1)}$ given $\mathbf{V}_{ij}^{(n)}$. To formulate the recursion, we write (22) more concisely as

$$\mathbf{P}\mathbf{V}_{ij}=\mathbf{R},\quad \mathbf{P}=\begin{pmatrix}a & b\\ c & d\end{pmatrix},\quad \mathbf{R}=\begin{pmatrix}q\\ r\end{pmatrix}, \tag{23}$$

where $\mathbf{P}$ and $\mathbf{R}$ depend on values of $\mathbf{V}$ at the neighboring points $(i\pm\alpha,\ j\pm\beta)$, with $\alpha,\beta=0,1$ and $\alpha+\beta\neq 0$.

The recursion or iteration is now simply

$$\mathbf{P}^{(n,n+1)}\mathbf{V}_{ij}^{(n+1)}=\mathbf{R}^{(n,n+1)}; \tag{23′}$$

the notation $\mathbf{P}^{(n,n+1)}$, $\mathbf{R}^{(n,n+1)}$ means that we sweep the mesh at each iterative step from left to right and from bottom to top, using the "new" values $\mathbf{V}_{i-\alpha,j-\beta}^{(n+1)}$ and the "old" values $\mathbf{V}_{i+\alpha,j+\beta}^{(n)}$. This procedure is sometimes *collective successive relaxation* (Brandt, 1977). It is a straightforward generalization of successive relaxation for a single Poisson equation. Of course, the situation in which $\det\mathbf{P}^{(n,n+1)}=0$ should be carefully avoided, but it was in fact not encountered in the test cases we studied.

Furthermore, in cases in which the problem was of mixed type, that is, hyperbolic subdomains were actually present, it turned out that it was important to use a *continuity method*, and start the iterations by solving a purely elliptic problem first. The simplest way to achieve that seemed to introduce instead of the parameter $f$ in (11) the quantity $f_n=f(1+\epsilon^n)$, so that $f^{(n)}>f$ for all $n$ and $f^{(n)}\downarrow f$ as $n$ increases. Using $\epsilon=0.5$ guaranteed in all our test cases that the first iteration corresponded to the solution of an elliptic system and that for the value of $n$ for which the iterative process was stopped, $f^{(n)}$ was sufficiently close to $f$ so as not to contribute significantly to discretization error.

The *convergence criterion* used to stop iteration was

$$\frac{1}{V_m}\max_{i,j}|\mathbf{V}_{ij}^{(n+1)}-\mathbf{V}_{ij}^{(n)}|<\eta,$$

where $V_m=\max_{i,j}|\mathbf{V}_{ij}|$ and $|\mathbf{W}|$ stands for the Euclidean length of the 2-vector $\mathbf{W}$, $\mathbf{W}=(w_1,w_2)$, i.e., $|\mathbf{W}|^2$

$=w_1^2+w_2^2$. It seems that $\eta=10^{-2}$ is quite sufficient for the intended application; using smaller $\eta$ did not improve the actual accuracy of the solutions very much, while necessitating a considerably larger number of iterations (Ghil and Shkoller, 1976).

*b. Test cases*

The most general two-dimensional velocity field can be written as

$$\mathbf{V}=\hat{\mathbf{k}}\times\nabla\psi+\nabla\chi, \tag{24}$$

where $\hat{\mathbf{k}}$ is the unit vector perpendicular to the plane of the flow, $\nabla$ is the two-dimensional gradient $[\nabla=(\partial/\partial_x,\partial/\partial_y)]$, $\psi$ a streamfunction and $\chi$ is a velocity potential. We study solutions of the diagnostic system (4) which can be represented by (24) with

$$\psi=-(V_0/k)\sin kx\sin ky,\quad \chi=(\tfrac{1}{2})(\lambda x^2+\mu y^2), \tag{25}$$

and $2\pi/k=L=6\times10^6$ m. These solutions include in particular the nondivergent velocity fields discussed by Paegle and Paegle (1974) for elliptic cases and by Ghil and Shkoller (1976) for mixed-type cases; the nondivergent solutions are characterized by $\lambda+\mu=0$.

Since (4) is nonlinear, its type will depend on the solution we seek. In other words, we shall be able to adjust the amount of hyperbolicity, as well as the divergence, by varying the parameters $V_0$, $\lambda$ and $\mu$.

Using (24) and (25), we obtain

$$\delta\equiv u_x+v_y=\lambda+\mu=\text{constant}, \tag{26a}$$

$$u_y+v_x=0, \tag{26b}$$

$$u_x-v_y=2kV_0\cos kx\cos ky+\lambda-\mu, \tag{26c}$$

$$\zeta\equiv v_x-u_y=2kV_0\sin kx\sin ky. \tag{26d}$$

To exhibit the dependence of the ellipticity criterion for (4) on $V_0$, $\lambda$ and $\mu$ as clearly as possible, we start with (9), which becomes, using (26b),

$$\begin{aligned}\mathfrak{D}&=(f+v_x-u_y)^2-(u_x-v_y)^2\\&=(f+v_x-u_y+u_x-v_y)(f+v_x-u_y-u_x+v_y).\end{aligned} \tag{9'}$$

Substitution of Eqs. (26c) and (26d) into (9'), and use of elementary trigonometric identities, yields

$$\mathfrak{D}=4k^2V_0^2\left[\frac{f+\gamma}{2kV_0}+\cos k(x-y)\right]\times\left[\frac{f-\gamma}{2kV_0}-\cos k(x+y)\right], \tag{27}$$

where we have introduced $\gamma$, defined as

$$\gamma=\lambda-\mu. \tag{27'}$$

From (27) it is evident that $\mathfrak{D}$ will be positive for all $(x,y)$ in our rectangular domain, i.e., for $0\leqslant kx,\ ky\leqslant 2\pi$, only if both the conditions

$$|\gamma|<f, \tag{28a}$$

$$V_0<\frac{f-|\gamma|}{2k}, \tag{28b}$$

are satisfied. We shall study only cases for which (28a) is satisfied; the ratio of $V_0$ to the critical velocity $V_*$, where

$$V_*=\frac{f-|\gamma|}{2k}, \tag{29}$$

is then a measure of the amount of hyperbolicity associated with a solution of form (24). Specifically, for $\gamma=0$,

$$V_*=\frac{f}{2k}\approx47.75\text{ m s}^{-1},$$

and this is the critical value corresponding in particular to those nondivergent solutions for which $\lambda=\mu=\lambda-\mu=0$.

*c. Numerical results*

To check the accuracy of the method described in Section 4a to solve system (11), we had to have a collection of exact solutions of (11). Substituting $\mathbf{V}$ given by (24) and (25) into the left-hand sides of (11), we obtain the corresponding right-hand sides of the system. The test of the method is then to reproduce as accurately as possible the known solution by using the information provided by the prescribed right-hand side and boundary conditions.

In the notation of (12), given $\mathbf{V}$, we obtain $g$, $h$ as known functions of $x$ and $y$:

$$G(\mathbf{V}'',\mathbf{V}')=g(x,y), \tag{12a'}$$

$$H(\mathbf{V}'',\mathbf{V}')=h(x,y). \tag{12b'}$$

Here $\mathbf{V}''$, $\mathbf{V}'$ are computed by analytic differentiation from $\mathbf{V}$ given by (24) and (25), with certain values of the parameters $V_0$, $\lambda$ and $\mu$. Then $G$ and $H$ are discretized as indicated in Section 4a and the discrete form of (12') is solved by the algorithm expressed in Eqs. (22) and (23'). We denote the result of this procedure by $\mathbf{V}_\Delta$.

The accuracy of the method can be evaluated by a suitable measure of the difference $\mathbf{V}_\Delta-\mathbf{V}$. This difference is the error of the numerical method. The most useful and widely used measures, or norms, are the rms error and the maximum (max) error.

We performed computations with two mesh sizes, corresponding to $M=M_1=25$ and $M=M_2=50$, i.e., $\Delta=\Delta_1=240$ km and $\Delta=\Delta_2=120$ km. Actually, we used $M_2=49$, because of programming considerations, but will refer in the sequel for simplicity to $M_2=50$. These mesh sizes are approximately equal to those currently

TABLE 1. Results of computations for System (11) with the algorithm given by (22) and (23′). The exact solutions are given by (24) and (25), and depend on the parameters $V_0$, $\lambda$ and $\mu$. The divergence $\delta$ of the analytical solution is given by (26a), as $\delta=\lambda+\mu$, and $\gamma$ is given by (27′), as $\gamma=\lambda-\mu$. The degree of hyperbolicity is expressed by the ratio of the maximum nondivergent velocity $V_0$ to the critical velocity $V_*$, where $V_*$ is given by (29), and by the relative extent of the subdomains of hyperbolicity $p$. The number of iterations $N$ required to satisfy our convergence criterion, with $\eta=10^{-2}$, is next. The subcolumns headed 10% refer to an initial guess which differs by a 10% random error from the exact solution; heading 20% refers to 20% random error in the initial guess, as explained in the text. $L_2(u)$ and $L_2(v)$ stand for the rms differences between the numerical and the analytical $u$ and $v$, respectively; these rms differences are expressed as percentages of the rms norms of $u$ and $v$ themselves. $\Delta u$ and $\Delta v$ are the actual maximum errors in $u$ and $v$ (m s$^{-1}$); they should be compared with $V_0$ in the first column. $L_2(\psi)$ and $L_\infty(\psi)$ give the rms error and maximum error in the streamfunction $\psi$, as computed from the numerical solution $(u,v)$ by solving the Poisson equation (30); they are also expressed as percentages of the rms norm and the maximum norm of $\psi$ itself. The results seem to indicate that, in spite of the severity of the test cases, a moderate number of iterations is required and reasonable accuracy is obtained. Furthermore, the results depend very little on the quality of the initial guess.

| $V_0$ (m s$^{-1}$) | $\delta$ ($\times 10^5$ s$^{-1}$) | $\gamma$ ($\times 10^5$ s$^{-1}$) | $V_*$ (m s$^{-1}$) | $p$ (%) | $N$ 10% | $N$ 20% | $L_2(u)$[%] 10% | $L_2(u)$[%] 20% | $L_2(v)$[%] 10% | $L_2(v)$[%] 20% | $\Delta u$[m s$^{-1}$] 10% | $\Delta u$[m s$^{-1}$] 20% | $\Delta v$[m s$^{-1}$] 10% | $\Delta v$[m s$^{-1}$] 20% | $L_2(\psi)$[%] 10% | $L_2(\psi)$[%] 20% | $L_\infty(\psi)$[%] 10% | $L_\infty(\psi)$[%] 20% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 55 | 0 | 0 | 47.75 | 21.92 | 24 | 20 | 5.65 | 5.72 | 5.68 | 5.79 | 5.60 | 5.78 | 5.44 | 5.29 | 4.84 | 5.15 | 10.45 | 10.74 |
| 55 | 2 | 0 | 47.75 | 21.92 | 25 | 21 | 3.49 | 3.52 | 3.44 | 3.50 | 5.49 | 5.65 | 5.46 | 5.36 | 4.79 | 5.08 | 10.40 | 10.67 |
| 55 | 5 | 1 | 42.97 | 28.96 | 36 | 27 | 1.49 | 1.43 | 2.17 | 2.12 | 6.21 | 5.83 | 6.32 | 5.99 | 4.68 | 4.84 | 10.52 | 10.53 |
| 54 | 0 | 0 | 47.75 | 21.92 | 22 | 19 | 5.20 | 5.35 | 5.23 | 5.41 | 5.06 | 5.30 | 4.89 | 4.78 | 4.61 | 4.95 | 9.74 | 10.10 |
| 54 | 2 | 0 | 47.75 | 21.92 | 23 | 20 | 3.17 | 3.24 | 3.13 | 3.23 | 4.95 | 5.17 | 4.92 | 4.86 | 4.56 | 4.87 | 9.58 | 10.02 |
| 54 | 5 | 1 | 42.97 | 23.20 | 29 | 24 | 1.30 | 1.28 | 1.89 | 1.90 | 4.89 | 4.93 | 5.10 | 4.99 | 4.41 | 4.62 | 9.51 | 9.75 |
| 52 | 0 | 0 | 47.75 | 21.92 | 18 | 16 | 4.58 | 4.81 | 4.62 | 4.89 | 4.19 | 4.49 | 3.97 | 3.98 | 4.32 | 4.70 | 8.64 | 9.05 |
| 52 | 2 | 0 | 47.75 | 21.92 | 19 | 17 | 2.71 | 2.83 | 2.69 | 2.83 | 4.08 | 4.37 | 4.01 | 4.02 | 4.25 | 4.61 | 8.57 | 8.97 |
| 52 | 5 | 1 | 42.97 | 20.96 | 24 | 21 | 1.08 | 1.08 | 1.55 | 1.61 | 3.92 | 4.09 | 4.11 | 4.09 | 4.02 | 4.27 | 8.34 | 8.61 |
| 50 | 0 | 0 | 47.75 | 21.92 | 17 | 15 | 4.12 | 4.43 | 4.17 | 4.50 | 3.55 | 3.90 | 3.36 | 3.45 | 4.03 | 4.47 | 7.78 | 8.28 |
| 50 | 2 | 0 | 47.75 | 21.92 | 17 | 15 | 2.39 | 2.56 | 2.38 | 2.56 | 3.45 | 3.79 | 3.38 | 3.51 | 4.01 | 4.42 | 7.74 | 8.20 |
| 50 | 5 | 1 | 42.97 | 20.96 | 20 | 18 | 0.91 | 0.94 | 1.35 | 1.42 | 3.24 | 3.49 | 3.38 | 3.50 | 3.79 | 4.07 | 7.45 | 7.78 |

used in numerical weather prediction on a global and a regional scale, respectively.

The values of $V_0$ used in the computations ranged up to $V_0=55$ m s$^{-1}$, which is very high for typical synoptic-scale winds. The values of $\lambda$ and $\mu$ used were such as to produce divergences up to $\delta=\lambda+\mu=5\times10^{-5}$ s$^{-1}$, which are also abnormally high for synoptic-scale wind fields [cf. Palmén and Newton (1969), Fig. 6.4a and Table 11.1, for instance].

As an initial guess $\mathbf{V}^{(0)}$ for our computations we used the correct solution $\mathbf{V}$ to which a random perturbation was added: at each grid point we added to both $u$ and $v$ a number obtained by sampling a random variable distributed uniformly between $-S$ and $+S$. The limit $S$ of the distribution was given in each case by $S=sV_0$. For $s$, we took the values $s=s_1=0.1$, or 10% of $V_0$, $s=s_2=0.2$, or 20% of $V_0$, and $s=s_3=0.3$, or 30% of $V_0$. Taking into account the large values of $V_0$ we used, it is clear that $S$ is quite large and exceeds considerably in all our examples the limitations on measurement error in wind included in GARP requirements (cf. Bengtsson, 1975, p. 3), which are 3 m s$^{-1}$ in mid-latitudes and 2 m s$^{-1}$ in the tropics. It is therefore reasonable to assume that an initial guess with comparable accuracy could be made in operational situations. We expect to address this question in further work on this problem.

The most important results of the computations are summarized in Tables 1 and 2. Some additional results will be commented upon with reference to those in the tables.

Both tables are organized as follows. The solution for which the results are given is identified by $V_0$, $\delta$ and $\gamma$, given in the first three columns in this order. Note that $\delta$ and $\gamma$ define uniquely $\lambda$ and $\mu$, and hence $\chi$, while $V_0$ defines $\psi$. The next two columns give the critical velocity $V_*$ and the number of points at which (4) was hyperbolic for the given solution $(V_0,\lambda,\mu)$, as a percentage $p$ of the total number of grid points; the ratio $V_0/V_*$ and the number of hyperbolic points are measures of the amount of hyperbolicity of the solution.

The other columns in both tables contain information on the numerical iteration and the results. Each column is divided into two subcolumns: the one to the left contains results for $M=25$ and $s=s_1$. Table 1 differs from Table 2 in the content of the subcolumns to the right: Table 1 has results for $M=25$ and $s=s_2$, while Table 2 has results for $M=50$ and $s=s_1$.

The result columns in the tables give $N$, the number of iterations to convergence, the rms errors in $u$ and $v$ in percentages of the corresponding measure of $u$ and $v$ themselves, the actual maximum error in $u$ and $v$, and the rms and max error in $\psi$, in percentages of its corresponding norms. Exact definitions of the rms and max norms, and of the relative error as a ratio of error norm to solution norm, are given in Ghil and Balgovind (1977).

The streamfunction $\psi$ was computed by solving the Poisson equation:

$$\nabla^2\psi=\zeta, \quad \zeta=v_x-u_y, \tag{30}$$

subject to the boundary condition $\psi=0$. The right-hand side $\zeta$ of Eq. (30) was computed at each grid point $(i,j)$ by centered finite differences from the numerical

solution $V_\Delta$ of (11); the equation was then solved numerically by the standard procedure of successive overrelaxation, i.e.,

$$\psi^*_{ij} = -\Delta^2\zeta_{ij} + (\tfrac{1}{4})(\psi^{n+1}_{i-1,j} + \psi^{n+1}_{i,j-1} + \psi^n_{i+1,j} + \psi^n_{i,j+1}), \quad (31a)$$

$$\psi^{n+1}_{ij} = \psi^n_{ij} + \omega(\psi^*_{ij} - \psi^n_{ij}). \quad (31b)$$

For $\omega = 1$ this reduces to ordinary successive relaxation (sometimes called Gauss-Seidel or Liebmann relaxation), but (31) produces faster convergence for $1 < \omega < 2$. We used $\omega = 1.8$, which is in many cases a good guess for the optimal value of $\omega$ (e.g., Forsythe and Wasow, 1960, Section 22).

Table 1 concentrates on the highest values of $V_0$ for which solutions were obtained, 50 m s$^{-1} \leqslant V_0 \leqslant$ 55 m s$^{-1}$. These values are indeed quite high compared to typical synoptic-scale values of 15–30 m s$^{-1}$, and it would be most unusual to obtain even higher values from observations. Divergence values in this table range up to $\delta = 5\times10^{-5}$ s$^{-1}$, which should be compared with typical synoptic-scale values of $1-3\times10^{-5}$ s$^{-1}$. Ratios of $V_0/V_*$ range up to 1.28 (for $V_0 = 55$ m s$^{-1}$, $V_* = 42.97$ m s$^{-1}$) and the percentage extent of hyperbolic subdomains ranges up to $p = 29.0\%$; both the values of $V_0/V_*$ and of $p$ show that the corresponding test cases were rather stringent tests of the ability of the numerical method to handle solutions with hyperbolic subdomains.

The errors in this table do not exceed 5.8% for the rms error. Max errors are naturally higher, but still less than 10.9% in the most extreme cases. This is a considerable reduction from the error in the initial guess, which was 20% in the max norm and ranged up to 21% in the rms norm.

We notice from this table that the magnitude of the error in the initial guess does not seem to affect significantly the results: errors for $s = s_2$ (right subcolumns) are only slightly larger than those for $s = s_1$ (left subcolumns). However, some preliminary experiments in which a *geostrophic* initial guess was used yielded much smaller errors in the computed solution, i.e., a gain in accuracy by a factor up to 2. This might be due simply to the smoothness of the geostrophic guess compared to the randomly perturbed initial guesses. We intend to study this matter further.

Table 2 concentrates on a comparison of numerical results with the two mesh sizes, corresponding to $M = 25$ and $M = 50$. Velocities of solutions tabulated here range up to 48 m s$^{-1}$, which is still very high, divergences again to $5\times10^{-5}$ s$^{-1}$, ratios $V_0/V_*$ to 1.44 (for $V_0 = 48$ m s$^{-1}$, $V_* = 33.42$ m s$^{-1}$), which is even higher than in Table 1, and the extent of hyperbolic subdomains to $p = 29.3\%$. Errors are considerably smaller than in Table 1, and they are much smaller for $M = 50$ than for $M = 25$. Typically for $M = 25$ the rms error is about 1–3%, and the max error is 4–7%. For $M = 50$ the rms error is typically 0.3–1.7% and the max error is 1.8–2.6%. The mesh sizes are still too large to conclude on the exact order of accuracy of the method; the results

TABLE 2. Results for the numerical solution of System (11), comparing computations with two different grid sizes. The column headings and contents are the same as in Table 1. The subcolumns headed 25 refer to $M = 25$, or a $25\times25$ grid, with grid size $\Delta = 240$ km; the heading 50 refers to $M = 50$, or a $50\times50$ grid, with $\Delta = 120$ km. A marked improvement in accuracy obtains with grid refinement, without an increase in the required number of iterations.

| $V_0$ (m s$^{-1}$) | $\delta$ ($\times10^5$ s$^{-1}$) | $\gamma$ ($\times10^5$ s$^{-1}$) | $V_*$ (m s$^{-1}$) | $p$ (%) | $N$ 25 | $N$ 50 | $L_2(u)$[%] 25 | $L_2(u)$[%] 50 | $L_2(v)$[%] 25 | $L_2(v)$[%] 50 | $\Delta u$[m s$^{-1}$] 25 | $\Delta u$[m s$^{-1}$] 50 | $\Delta v$[m s$^{-1}$] 25 | $\Delta v$[m s$^{-1}$] 50 | $L_2(\psi)$[%] 25 | $L_2(\psi)$[%] 50 | $L_\infty(\psi)$[%] 25 | $L_\infty(\psi)$[%] 50 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 48 | 0 | 0 | 47.75 | 7.84 | 18 | 11 | 3.67 | 1.38 | 3.71 | 1.79 | 3.00 | 0.94 | 2.86 | 1.25 | 3.67 | 1.52 | 6.96 | 2.23 |
| 48 | 2 | 0 | 47.75 | 7.84 | 18 | 11 | 2.08 | 0.79 | 2.06 | 1.00 | 2.91 | 0.91 | 2.89 | 1.23 | 3.66 | 1.50 | 6.94 | 2.20 |
| 48 | 5 | 1 | 42.97 | 12.64 | 18 | 12 | 0.80 | 0.30 | 1.20 | 0.56 | 2.75 | 1.07 | 2.94 | 1.34 | 3.60 | 1.41 | 6.80 | 2.03 |
| 48 | 5 | 3 | 33.42 | 29.28 | 35 | | 0.74 | | 2.51 | | 4.06 | | 4.21 | | 3.53 | | 7.57 | |
| 47 | 0 | 0 | 47.75 | 0.00 | 18 | 12 | 3.49 | 1.33 | 3.53 | 1.73 | 2.76 | 0.89 | 2.66 | 1.16 | 3.53 | 1.48 | 6.62 | 2.17 |
| 47 | 2 | 0 | 47.75 | 0.00 | 18 | 12 | 1.95 | 0.75 | 1.94 | 0.95 | 2.68 | 0.86 | 2.69 | 1.14 | 3.53 | 1.47 | 6.61 | 2.14 |
| 47 | 5 | 1 | 42.97 | 12.64 | 17 | 12 | 0.76 | 0.29 | 1.14 | 0.54 | 2.56 | 0.89 | 2.77 | 1.24 | 3.53 | 1.40 | 6.54 | 1.98 |
| 47 | 5 | 3 | 33.42 | 20.96 | 24 | | 0.65 | | 2.25 | | 3.21 | | 3.39 | | 3.64 | | 7.04 | |
| 46 | 0 | 0 | 47.75 | 0.00 | 19 | 12 | 3.25 | 1.31 | 3.29 | 1.72 | 2.51 | 0.87 | 2.42 | 1.12 | 3.33 | 1.47 | 6.20 | 2.12 |
| 46 | 2 | 0 | 47.75 | 0.00 | 18 | 12 | 1.83 | 0.73 | 1.82 | 0.93 | 2.47 | 0.84 | 2.50 | 1.11 | 3.40 | 1.45 | 6.30 | 2.09 |
| 46 | 5 | 1 | 42.97 | 12.64 | 17 | 12 | 0.72 | 0.28 | 1.07 | 0.52 | 2.39 | 0.86 | 2.59 | 1.16 | 3.42 | 1.39 | 6.27 | 1.94 |
| 46 | 5 | 3 | 33.42 | 20.96 | 21 | | 0.60 | | 2.12 | | 2.89 | | 3.04 | | 3.48 | | 6.67 | |
| 44 | 0 | 0 | 47.75 | 0.00 | 19 | 12 | 2.96 | 1.28 | 2.99 | 1.69 | 2.14 | 0.84 | 2.09 | 1.65 | 3.09 | 1.45 | 5.53 | 2.03 |
| 44 | 2 | 0 | 47.75 | 0.00 | 19 | 12 | 1.58 | 0.69 | 1.57 | 0.88 | 2.09 | 0.81 | 2.13 | 1.04 | 3.10 | 1.43 | 5.64 | 2.00 |
| 44 | 5 | 1 | 42.97 | 4.32 | 18 | 12 | 0.62 | 0.28 | 0.93 | 0.49 | 2.08 | 0.80 | 2.24 | 1.02 | 3.15 | 1.37 | 5.70 | 1.87 |
| 44 | 5 | 3 | 33.42 | 20.96 | 16 | | 0.53 | | 1.91 | | 2.45 | | 2.61 | | 3.35 | | 6.12 | |
| 40 | 0 | 0 | 47.75 | 0.00 | 19 | 12 | 2.50 | 1.24 | 2.52 | 1.64 | 1.65 | 0.77 | 1.56 | 0.93 | 2.71 | 1.40 | 4.72 | 1.88 |
| 40 | 2 | 0 | 47.75 | 0.00 | 19 | 12 | 1.24 | 0.62 | 1.24 | 0.80 | 1.60 | 0.74 | 1.60 | 0.93 | 2.72 | 1.39 | 4.74 | 1.86 |
| 40 | 5 | 1 | 42.97 | 0.00 | 18 | 12 | 0.49 | 0.23 | 0.73 | 0.44 | 1.63 | 0.72 | 1.73 | 0.90 | 2.79 | 1.34 | 4.87 | 1.74 |
| 40 | 5 | 3 | 33.42 | 20.96 | 18 | 12 | 0.41 | 0.18 | 1.50 | 0.88 | 1.86 | 0.76 | 1.98 | 1.14 | 2.94 | 1.34 | 5.13 | 1.77 |
| 36 | 0 | 0 | 47.75 | 0.00 | 19 | 12 | 2.16 | 1.21 | 2.18 | 1.60 | 1.27 | 0.70 | 1.18 | 0.83 | 2.41 | 1.37 | 4.04 | 1.75 |
| 36 | 2 | 0 | 47.75 | 0.00 | 19 | 12 | 0.99 | 0.56 | 0.99 | 0.72 | 1.24 | 0.68 | 1.22 | 0.82 | 2.42 | 1.36 | 4.07 | 1.74 |
| 36 | 5 | 1 | 42.97 | 0.00 | 18 | 11 | 0.39 | 0.21 | 0.59 | 0.40 | 1.28 | 0.70 | 1.34 | 0.84 | 2.51 | 1.33 | 4.23 | 1.68 |
| 36 | 5 | 3 | 33.42 | 12.64 | 18 | 11 | 0.32 | 0.16 | 1.20 | 0.78 | 1.45 | 0.73 | 1.52 | 0.88 | 2.62 | 1.33 | 4.38 | 1.73 |

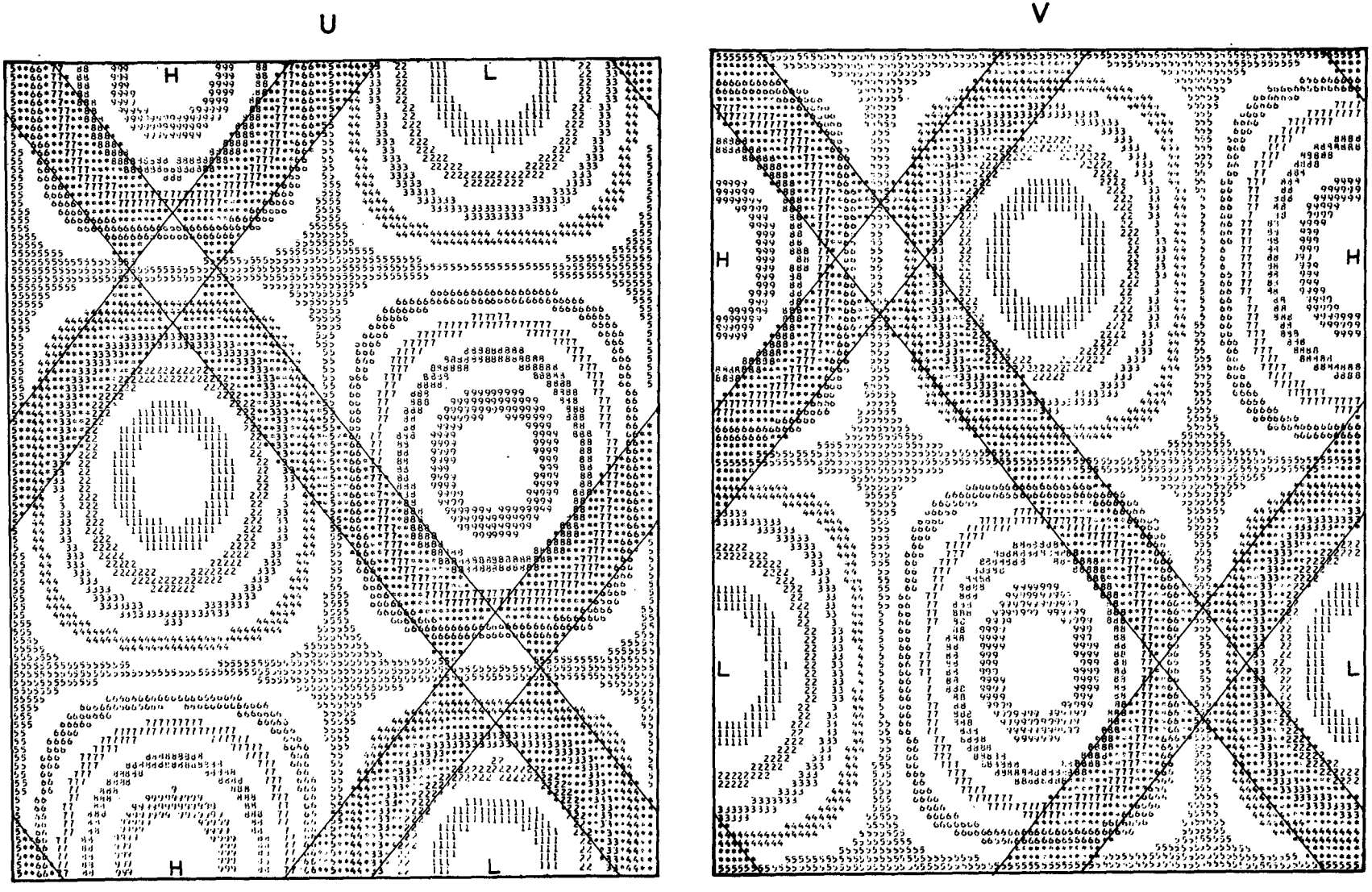


FIG. 1. The numerical solution of (11), showing the $u$ field (a) and $v$ field (b), for a test case with analytical solution given by (24) and (25), with $V_0=54$ m s$^{-1}$, $\lambda=\mu=0$; it corresponds to $V_0/V_*=1.13$ and to $p$ grid points, $p=22\%$ being hyperbolic. The plotting was done by subdividing the range of the plotted variable into 17 equal intervals and assigning digits from 1 to 9 in increasing order to each second interval in succession; intervals marked by blanks separate those marked by digits. Half-intervals of blanks are assigned to the very beginning and the very end of the range, thus: b11bb22bb33bb...88bb99b. The grid points where the minimum and the maximum of the plotted variable obtained are marked by an L and an H, respectively. The hyperbolic subdomains appear as darker areas: the grid points at which the analytic solution is hyperbolic are marked by asterisks if they fall within an interval of blanks, and by a darker digit otherwise. Elliptic and hyperbolic regions are separated by thin lines. The numerical solution is visually almost indistinguishable from the exact one. It was obtained on a mesh of 25×25 points ($\Delta x=\Delta y=240$ km) and interpolated linearly to 50×50 points.

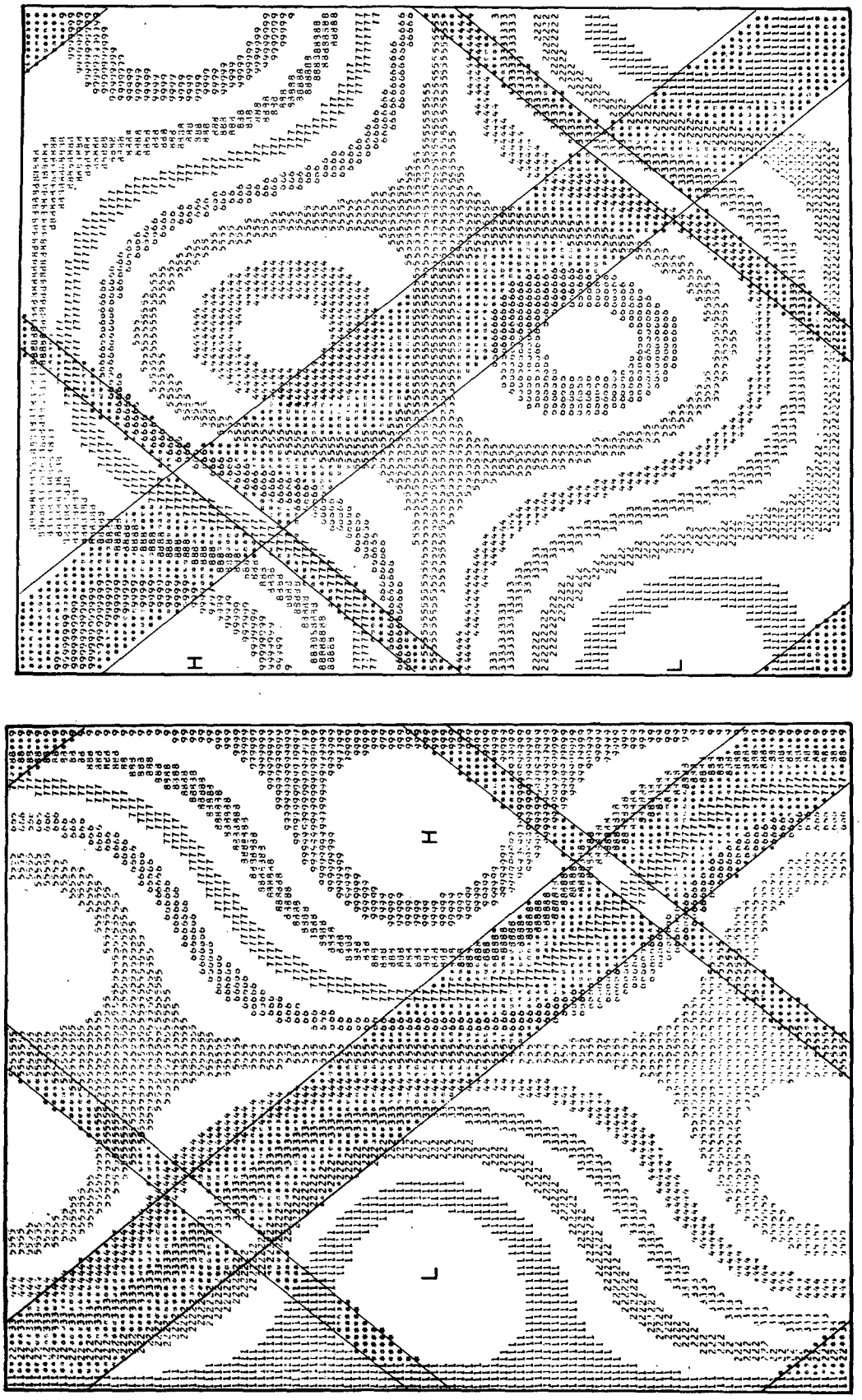


FIG. 2. As in Fig. 1, but for $V_0 = 54$ m s$^{-1}$, $\lambda = 3\times10^{-5}$ s$^{-1}$, $\mu = 2\times10^{-5}$ s$^{-1}$; this test case has $V_0/V_* = 1.26$, $\delta = 5\times10^{-5}$ s$^{-1}$ and $\rho = 23\%$.

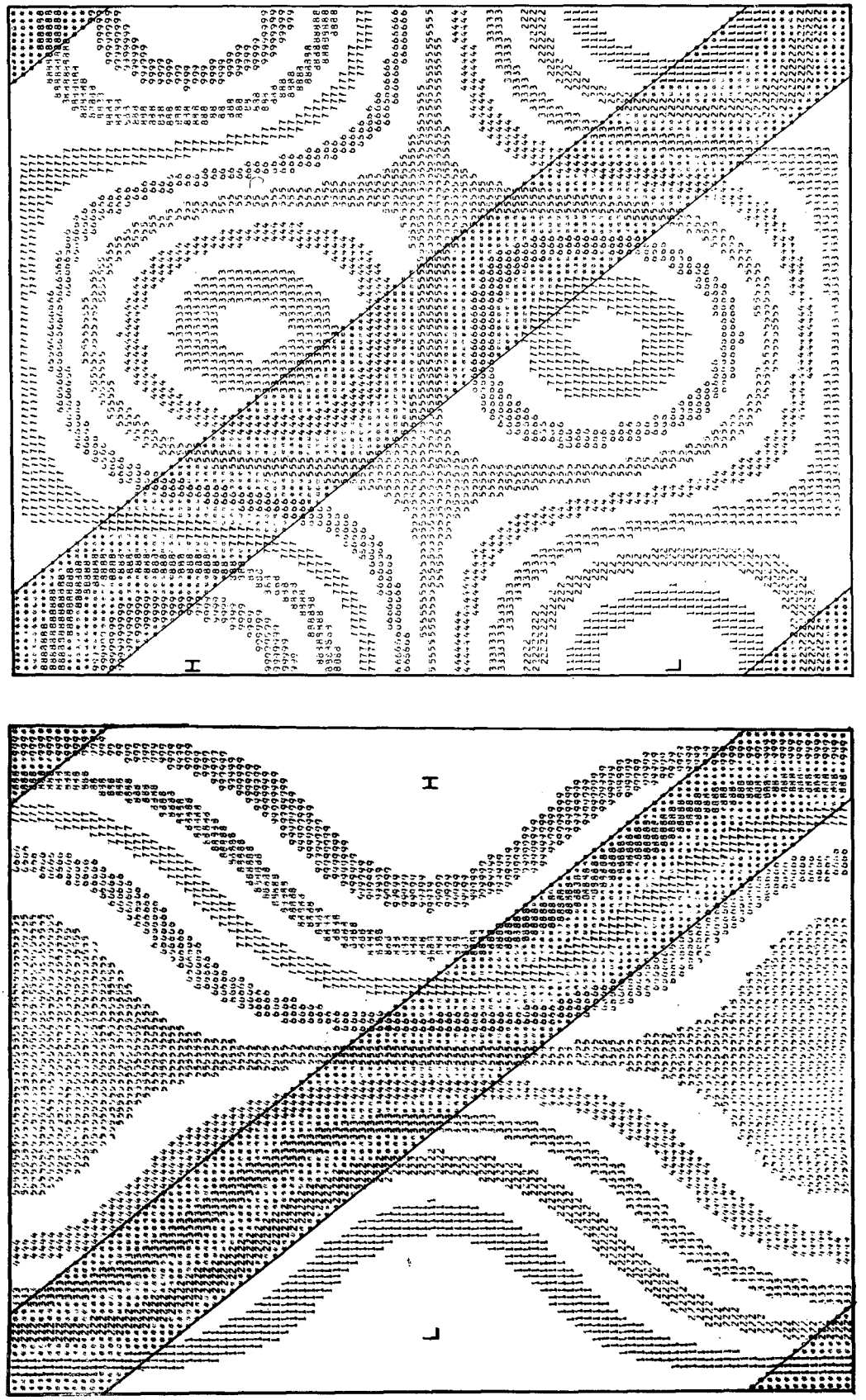


FIG. 3. As in Fig. 1 but for $V_0 = 47$ m s$^{-1}$, $\lambda = 4 \times 10^{-5}$ s$^{-1}$ and $\mu = 10^{-6}$ s$^{-1}$; this test case has $V_0/V_* = 1.41$, $\delta = 5 \times 10^{-5}$ s$^{-1}$ and $\phi = 21\%$.

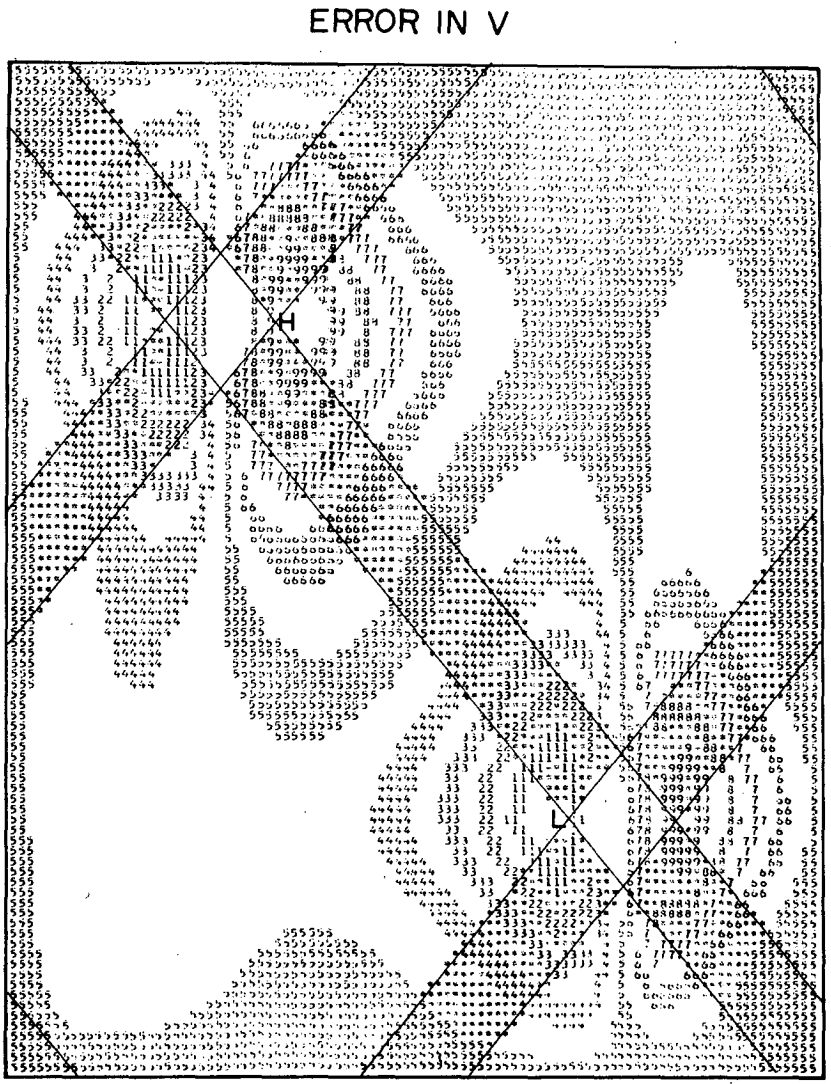


FIG. 4. The error in $v$ for the solution plotted in Fig. 1. The maximum error is about 5 m s$^{-1}$, the rms error is about half of that, or 2.8 m s$^{-1}$; these have to be compared with a maximum velocity of $V_0=54$ m s$^{-1}$. Errors are largest near the intersection of the bands of hyperbolicity and decrease rapidly away from these intersections. The error in the interior of the elliptic subdomains is practically zero.

seem to suggest that it is indeed second-order accurate, as one might be led to expect from the discretization chosen.

We notice that for $M=50$ the numerical iteration did not converge in some cases in which the computation with $M=25$ produced convergence. This phenomenon was observed also when initial guesses with $s=s_2$ and $s=s_3$ were used on the finer grid. Some cases with larger initial errors failed to converge even on the coarser grid, while convergence was obtained when $s=s_1$. Work on eliminating these difficulties is in progress.

The number of iterations $N$ never exceeds 36, and is typically 10–20 in all reasonable cases, with $V_0$ and $\delta$ not too large. This compares quite favorably with updating methods. Indeed, one iteration step of the present method for (11) requires about the same number of operations as a time step in the numerical integration of the prognostic system (1). Time steps of not more than 5–10 min can be taken for mesh sizes comparable to the ones we used when integrating (1). About 300 time steps are thus required for dynamic initialization by updating of (1) under the most favorable circumstances; as noted already in Section 1, the results for an updating process of this duration are still not satisfactory. On the other hand, the error in our results can be reduced even further by using a more stringent convergence criterion of $\eta=10^{-3}$ instead of $\eta=10^{-2}$ as reported here; this will only cause an increase of $N$ to about 100 (Ghil and Shkoller, 1976).

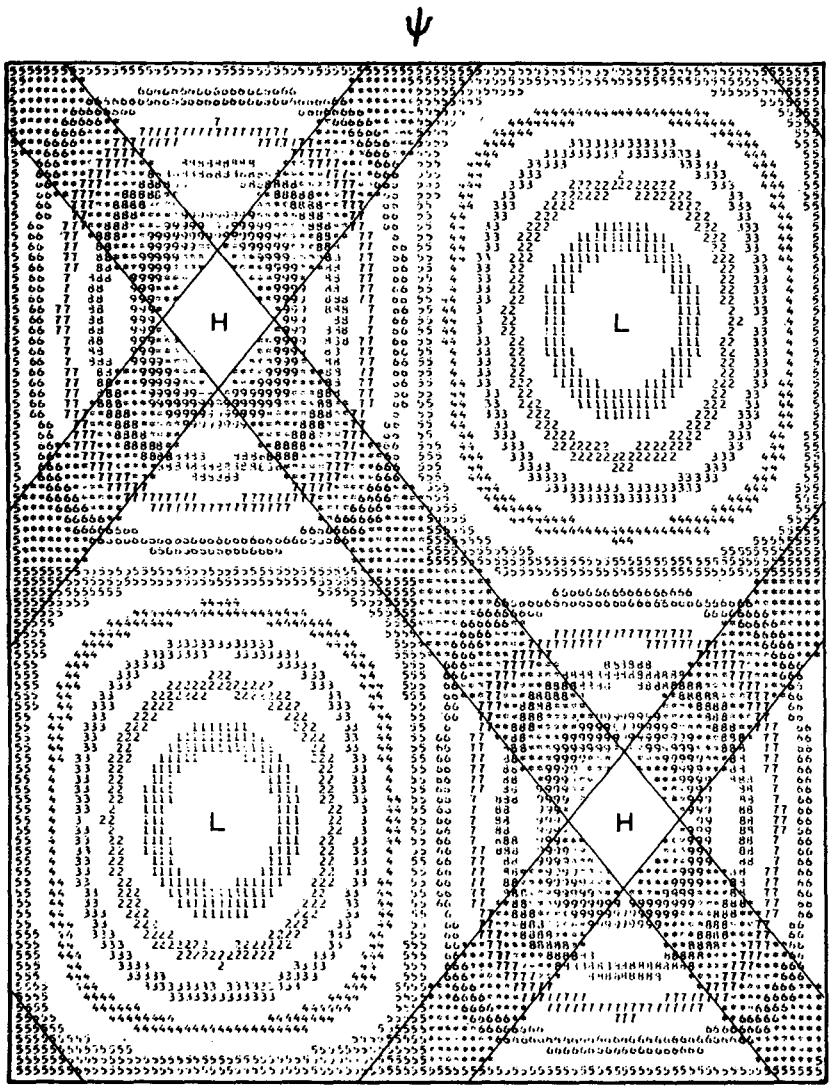


FIG. 5. The streamfunction $\psi$ for the solutions plotted in Fig. 1 and in Fig. 2. The remarks made about the visual accuracy of $u$ and $v$ apply to $\psi$ as well. This indicates that the method proposed here is also quite appropriate for solving the classical balance equation in mixed-type cases.

We notice in both tables that the errors in $\psi$ are never much larger than and mostly quite comparable to those in $u$ and $v$, and this in spite of the fact that (30) was not solved in the most accurate manner possible. One could for instance evaluate the right-hand side $\zeta$ more accurately, and use (31) or other, faster methods (cf. Ghil and Balgovind, 1977) on a much finer grid. Even with the present solution methods, given by (23′) and (31), we conclude that the solution of system (11) with $d=0$, followed by the solution of Eq. (30), is an efficient way of solving the classical balance equation in mixed-type cases (also compare Ghil and Shkoller, 1976).

We further illustrate our results in Figs. 1–5. The darker areas represent *hyperbolic subdomains*. Figs. 1–3 give plots of $u$ and $v$ for three selected solutions: Fig. 1 gives the plots of the numerical solution corresponding to $V_0=54$ m s$^{-1}$, $\lambda=\mu=0$; Fig. 2 gives the plots for $V_0=54$ m s$^{-1}$, $\lambda=3\times10^{-5}$ s$^{-1}$, $\mu=2\times10^{-5}$ s$^{-1}$; and Fig. 3 gives the plots for $V_0=47$ m s$^{-1}$, $\lambda=4\times10^{-5}$ s$^{-1}$, $\mu=1\times10^{-5}$ s$^{-1}$. Figs. 1a, 2a and 3a show $u$, while Figs. 1b, 2b and 3b show the corresponding $v$. All the plotted solutions were obtained with the coarser mesh, i.e., $M=25$. We notice that the plotted solutions have some of the highest numerical errors in the tables and strong hyperbolicity, while two of them also have very large divergence. In spite of all this the plots are visually almost indistinguishable from those of the exact solutions; this can be seen here from the smoothness and symmetry of the plot. Including plots of the exact solu-

tions did not seem to serve a useful purpose, since they would only duplicate the ones of the numerical solutions.

Fig. 4 gives an actual error plot, at a suitably magnified scale. It indicates that in this case, as well as in all the others, the errors were mostly concentrated in the areas where more than one elliptic and one hyperbolic region were in close vicinity of one another, or, in other words, near the intersection of lines of parabolicity. This can be expected from the mathematical theory of mixed-type problems (Bitsadze, 1964, Chap. 4). The errors in the interior of elliptic subdomains were entirely negligible in all cases. It might be worthwhile to notice at this point that the structure of the subdomains of hyperbolicity in the cases we present, as well as the extent of these subdomains, could have been expected to present greater difficulties than those encountered in the numerical solution of mixed-type problems currently solved in transonic gas dynamics (e.g., Jameson, 1976).

Fig. 5 shows the numerical solution for $\psi$, in the case $V_0=54$ m s$^{-1}$. Notice that $\psi$ is independent of $\lambda$ and $\mu$, and its computation makes sense even when $d\neq 0$. The same comments as for Figs. 1–3 apply with regard to the visual accuracy of $\psi$.

This concludes the presentation of our numerical results. It seems appropriate to comment at this point on some directions of development in the numerical techniques, while keeping comments of a more general nature for the last section.

We expect that results can be further improved by one or more of the following modifications of the method: the use of smooth, perhaps geostrophic initial guesses; using the results of the computation on a coarse mesh as an initial guess for the iteration on a finer mesh (Brandt, 1977); refining the continuity method idea, for instance by other choices of the sequence $f^{(n)}$; and exploring different ideas to accelerate convergence of the iteration, so that higher accuracy can be obtained in the same computer time. The last point can certainly be implemented with ease in the computation of $\psi$ from $u$ and $v$. Studies on the contribution of discretization and roundoff to the total errors in $u$ and $v$ showed that round-off errors are a large part of the total error (Ghil and Shkoller, 1976). This clearly indicates that an improvement in the method for solving the discrete system (22) will considerably increase the accuracy of the numerical solution $\mathbf{V}_\Delta$ of (11).

After perfecting the numerical method, it will be necessary to study the effect that errors in the boundary data for $u$ and $v$ and errors in the geopotential measurements entering into the right-hand side of system (4) have on the accuracy of the numerical solution. Such a study has been carried out analytically for a linearized version of (4) in Ghil (1975b). After gaining some numerical experience, similar theoretical estimates could perhaps be given for the effect of errors in the data of (4) on its solution.

## 5. Concluding remarks

We have derived a diagnostic balance system compatible with the nonlinear shallow-fluid equations. A diagnostic system compatible with a baroclinic prognostic model governed by nonlinear primitive equations was derived in Ghil (1975b).

These diagnostic systems give the horizontal velocity components $u$, $v$ in terms of the mass field and of its first and second time derivatives. They are derived from the corresponding prognostic systems without recourse to any physical assumptions in addition to those used in formulating the prognostic systems themselves.

Equipped with the information provided by these diagnostic systems, we now wish to address the following important question: does knowledge of the mass field and of its time history fully determine the velocity field in an atmosphere governed by nonlinear primitive equations? The form of our diagnostic systems allows us first to define the knowledge required about the time history of the mass field precisely, as knowledge of the first two time derivatives of the mass field. Second, the exact derivation of the diagnostic systems from the prognostic ones allows us to answer the question in the affirmative: yes, the mass field and its first two time derivatives do determine the wind field, in the precise sense we defined, and in all generality. We only expect, however, the diagnostic systems we derive to be of practical use in mid-latitudes where measurements of the mass field are sufficiently accurate.

Considering further the question of the practicality of the present approach of *initialization by compatible balancing*, we notice first that the numerical solution of our barotropic diagnostic system was shown to be more efficient in terms of computing time than updating the corresponding prognostic system; it was also shown that the results of our method are sufficiently accurate. The numerical difficulties connected with the nonlinear mixed elliptic-hyperbolic type of the diagnostic system have been surmounted in cases in which the hyperbolic subdomains covered up to a third of the entire computational domain. Such cases seem highly unlikely in actual synoptic situations; therefore the numerical method we present should be useful for all the synoptic situations likely to be encountered in practical applications of the diagnostic system presented here. The numerical method formulated here or modified forms of it could also be tried for other equations of mixed type occurring in dynamic meteorology, such as the $\omega$ equation; the ellipticity of the latter equation depends on the sign of the vertical static stability.

We have given an interpretation of the ellipticity condition for our diagnostic system and, by implication, for the classical balance equation, in terms of the difference of the square of absolute vorticity and the square of total deformation. This interpretation could serve in the study of synoptic diagnoses. It could also help

to elucidate the reason for the rarity and limited extent of hyperbolic areas in real synoptic situations.

We turn now to questions related to the implementation of initialization by compatible balancing in an operational context. First we notice that the diagnostic system used would have to be compatible in finite difference form, rather than in differential form, with the prognostic system of the numerical weather prediction (NWP) model. This is not so difficult to achieve (Ghil, 1975b). Certainly more work needs to be done on the numerical solution of finite-difference diagnostic systems compatible in this sense with baroclinic prognostic systems of primitive equations.

Second, the application of this approach would seem most useful in limited areas of sparse velocity data, such as the oceans. Boundary data for the velocity are available along the continental coasts from conventional observing stations; along a southern tropical boundary in the North Pacific, say, velocity data can be obtained from cloud tracking by geostationary satellites (Bengtsson, 1975).

It remains to be seen how accurately the mass field itself and its first and second time differences can be determined from temperature data based on radiance measurements from polar orbiting satellites. Methods for the four-dimensional (4-D) assimilation of satellite-derived temperatures obtained during a period of 12 to 24 h preceding initialization time are continuously being refined; we expect these methods to produce in the near future sufficiently accurate fields of geopotential heights and of time differences of these heights. The 4-D analysis should provide adequate filtering of observational errors. Statistical and variational methods for 4-D analysis are currently being developed with this objective in mind (Ghil *et al.*, 1976, Ghil and Mosebach, 1977).

A balanced diagnostic system compatible with the prognostic system of the NWP model would seem ideally suited as a nonholonomic constraint in a variational formulation (cf. Stephens, 1970). In such a formulation the time derivatives of the mass field would be minimized, along with the difference between initial state and observations.

In any case, the use of compatible balancing appears to hold some promise that the effort spent in determining an initial state which is as accurate as possible will not be wasted by the prognostic model's "immune reaction" to the initial state thus determined. Initialization by compatible balancing seems feasible and should prevent initialization shock and the rapid deterioration of model forecasts through it.

*Acknowledgments.* One of the authors (M. G.) benefited greatly from conversations with Achi Brandt, Milton Halem, Eugene Isaacson, Carlos Leiva and Arne Sundström. His research was supported by a NAS/NRC Research Associateship at the Goddard Institute for Space Studies, and by NASA Grant NSG-5130 at the Courant Institute.

It is a pleasure to acknowledge helpful suggestions from the referees.

## REFERENCES

Bengtsson, L., 1975: Four-dimensional assimilation of meteorological observations. GARP Publ. Ser., No. 15, WMO-ICSU, Geneva, WMO, 76 pp.

Bitsadze, A. V., 1964: *Equations of the Mixed Type.* Macmillan, 160 pp.

Blumen, W., 1976a: Experiments in atmospheric predictability: Part I. Initialization. *J. Atmos. Sci.*, **33**, 161–169.

——, 1976b: On dynamical/statistical initialization for numerical weather prediction. *J. Atmos. Sci.*, **33**, 2338–2349.

Brandt, A., 1977: Multi-level adaptive solutions to boundary value problems. *Math. Comput.* (in press).

Charney, J. G., M. Halem and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160–1163.

Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics*, Vol. 2. Wiley-Interscience, 830 pp.

Forsythe, G. E., and W. R. Wasow, 1960: *Finite-Difference Methods for Partial Differential Equations.* Wiley, 444 pp.

Fjørtoft, R., 1962: A numerical method of solving certain partial differential equations of second order. *Geofys. Publ.*, **24**, 229–239.

Ghil, M., 1973: On balance and initialization. Rep. IMM-400, Courant Institute of Mathematical Sciences, New York University, 32 pp. [NTIS AD 772501]

——, 1975a: The initialization problem in numerical weather prediction. *Improperly Posed Boundary Value Problems*, A. Carasso and A. P. Stone, Eds., Pitman Publishing, 105–123.

——, 1975b: Initialization by compatible balancing. Rep. 75-16, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Va, 38 pp. [Also available from the author, Courant Institute, NYU].

——, and B. Shkoller, 1976: Wind laws for shockless initialization. *Ann. Meteor.* (Neue Folge), **11**, 112–115.

——, and R. Balgovind, 1977: A fast Cauchy-Riemann solver with nonlinear applications. Submitted to *Math. Comput.* [Available from M. G., Courant Institute, NYU]

——, and R. Mosebach, 1977: Asynoptic variational method. The GISS Sounding Temperature Impact Test Project 1976, Final Rep., Chap. 3, Analysis and Assimilation, M. Ghil and R. Dilling, Eds., NASA Institute for Space Studies, Goddard Space Flight Center, 3.32–3.49.

——, R. Dilling and H. Carus, 1976: A time-continuous statistical assimilation method for satellite-derived temperatures. *NASA Weather and Climate Program, Basic Science Review*, NASA Goddard Space Flight Center, Greenbelt, Md., 42–47.

Hinkelmann, K., 1951: Der Mechanismus des meteorologischen Lärmes. *Tellus*, **3**, 285–296.

Jameson, A., 1976: Numerical solution of nonlinear partial differential equations of mixed type. *Numerical Solution of Partial Differential Equations*, Vol. 3, B. Hubbard, Ed., Academic Press, 275–320.

Miyakoda, K., and R. W. Moyer, 1968: A method of initialization for dynamical weather forecasting, *Tellus*, **20**, 115–128.

Morel, P., G. Lefevre and G. Rabreau, 1971: On initialization and nonsynoptic data assimilation. *Tellus*, **28**, 197–206.

Nitta, T., and J. B. Hovermale, 1969: A technique of objective analysis and initialization for the primitive forecast equations. *Mon. Wea. Rev.*, **97**, 652–658.

Paegle, J., and J. N. Paegle, 1974: An efficient and accurate approximation to the balance wind with application to non-elliptic data. *Mon. Wea. Rev.*, **102**, 838-846.

Palmén, E., and C. W. Newton, 1969: *Atmospheric Circulation Systems, Their Structure and Physical Interpretation.* Academic Press, 603 pp.

Petterssen, S., 1953: On the relation between vorticity, deformation and divergence and the configuration of the pressure field, *Tellus*, **3**, 231–237.

Stephens, J. J., 1970: Variational initialization with the balance equation. *J. Appl. Meteor.*, **9**, 732–739.

Williamson, D. L., and A. Kasahara, 1971: Adaptation of meteorological variables forced by updating. *J. Atmos. Sci.*, **28**, 1313–1324.